REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-04-

0015

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	12/21/03	Final Report

4. TITLE AND SUBTITLE
Non-invasive techniques for monitoring human fatigue

5. FUNDING NUMBERS
F49620-00-1-0243

6. AUTHOR(S)
Qiang Ji

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Nevada at Reno

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dr. Larkin Willard
Air Force Office of Scientific Research,
4015 Wilson Blvd. Arlington, VA
22203-1954

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
unrestricted

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

In this report, we summarize our efforts in developing real time non-intrusive technology for monitoring human fatigue. Through this research, we have developed state of the art technologies and a prototype fatigue monitor for real time non-intrusive human fatigue monitoring. Our contributions include: 1) the development of various computer vision techniques for real-time and non-intrusive extraction of multiple fatigue parameters related to eyelid movements, gaze, head movement, and facial expressions, 2) the development of a probabilistic framework based on the Bayesian networks to model and integrate contextual and visual cues information for robust and accurate fatigue detection, and 3) systematic and scientific validation of the fatigue monitor. Experimental validation of our techniques using human subjects demonstrates the good measurement accuracy of our techniques. In addition, the validation also verifies the validity of the proposed fatigue parameters as well as that of the composite fatigue index computed by our fatigue monitor.

14. SUBJECT TERMS
Real time human fatigue monitoring

15. NUMBER OF PAGES
30

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

NON-INVASIVE TECHNIQUES FOR MONITORING HUMAN FATIGUE

Qiang Ji

Department of Computer Science
University of Nevada Reno
Department of Electrical, Computer, and System Eng.
Rensselaer Polytechnic Institute
jiq@rpi.edu

Final Report for AFOSR Project
F49620-00-1-0243

December, 2003

1. Summary

This is the final report for our AFOSR sponsored project: non-invasive techniques for monitoring human fatigue. Through this project, we develop a real time non-intrusive prototype human fatigue monitor. The fatigue monitor was subsequently validated by a study involving human subjects to correlate its output with a vigilance task performance. In this report, we first summarize our technical accomplishments, followed by a discussion of transitions related to this project. The latest paper reprints, publications, demos, and media coverage about this project may be found at http://www.ecse.rpi.edu/~qji/Fatigue/fatigue.html

2. Introduction

As combat systems become more and more sophisticated and reliable, the major limiting factor for operational dominance in a conflict is the warfighter. Eliminating the potential for fatigue while maintaining a high level cognitive and physical performance of the warfighter will create a fundamental change in war fighting and force employment. Developing a technology to detect and predict the degradation of psychomotor performance of a warfighter due to fatigue is therefore critical to ensure the success of a mission.

Many efforts [3,5,6,7,13,14,16,18] have been reported in the literature for developing active fatigue monitoring systems. Among different techniques, the best detection accuracy is achieved with techniques that measure physiological conditions like brain waves, heart rate, and pulse rate [15,18]. These techniques, while accurate, are limited to in-house study and are not applicable to many real world applications due to extremely intrusiveness.

People in fatigue exhibit certain visual behaviors easily observable from changes in facial features such as the eyes, head, and face. Visual behaviors that typically reflect a person's level of fatigue include eyelid movement, gaze, head movement, and facial expression. To make use of these visual cues, another increasingly popular and non-invasive approach assessing fatigue is through the analysis of one's video image using state-of-the-art technologies in computer vision. Techniques using computer vision are aimed at extracting and analyzing visual characteristics typically reflecting an operator's level of fatigue from the video images of the operator. Typical visual characteristics observable from the image of a person with reduced alertness level include slow eyelid movement, smaller degree of eye openness (or even closed), frequent nodding, yawning, narrowness in the line of sight, sluggish in facial expression, and sagging posture. The main advantage of this approach is that it is non-intrusive and therefore will receive a complete cooperation from the operator. In fact, a recent workshop [2], sponsored by the Department of Transportation (DOT) on driver's vigilance, concluded that computer vision represents the most promising non-invasive technology to monitor driver's vigilance.

Numerous efforts have been reported in the literature on developing active real-time image-based fatigue monitoring systems [3,5,7,14,16,18,42,43,44,45,46]. These efforts are primarily focused on detecting driver fatigue. Despite these efforts, real time non-intrusive human fatigue monitoring remains a largely unresolved issue. One deficiency with the current efforts is that they tend to use only a single visual parameter such as PERCLOS. Due to the inherent ambiguity and uncertainty associated with a single parameter, these systems tend to be less robust and accurate. To overcome this limitation, we propose to simultaneously use multiple fatigue parameters. All

these parameters, however imperfect they are individually, if combined systematically, can provide an accurate and robust characterization of a person's level of vigilance.

The work for this project consists of two major parts. The first part focuses on developing real time computer vision algorithms to compute various parameters to characterize eyelid movement, gaze movement, head movement, and facial expression. The second part focuses on building a probabilistic framework to model fatigue and to systematically combine different fatigue parameters, along with the relevant contextual information, to produce a composite fatigue index. Figure 1 summarizes our approach.

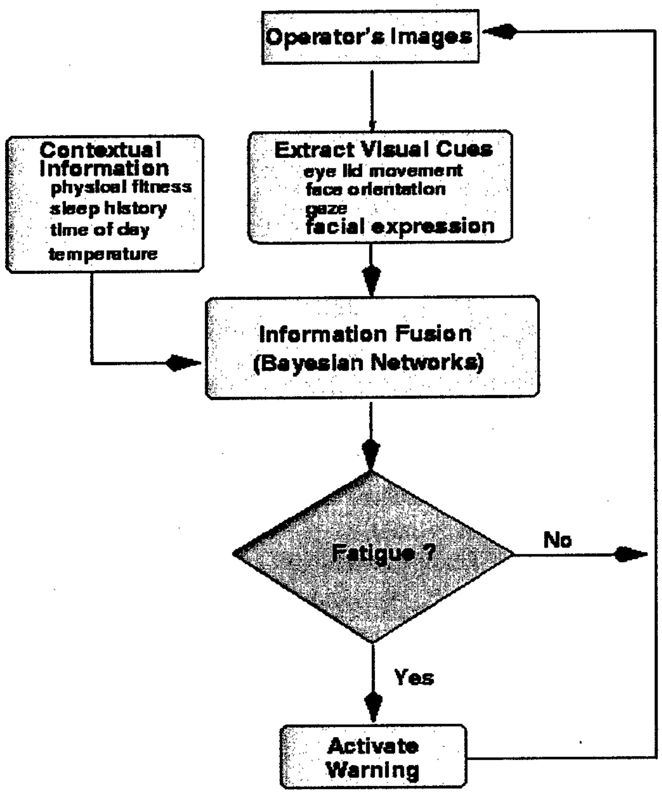


Figure 1: Flowchart of the proposed fatigue monitoring system

We have made significant progress in each of these two areas as detailed below.

3. Visual Cues Extraction

To monitor fatigue, we propose to monitor the subject's facial behaviors, identify visual cues typically characterizing a person's state of alertness, and develop computer vision algorithms to compute them non-intrusively in real time. In this section, we summarize the prototype computer vision system we have developed to achieve this goal. Details of the algorithms may be found in these publications [9-12,19,30-38].

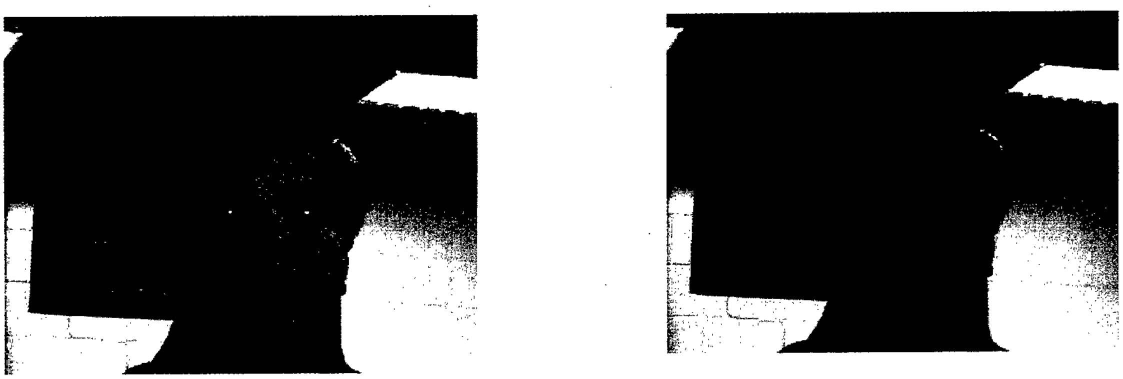
3.1 Hardware Setup

The main hardware components of our system consist of a remotely located CCD camera, a specially-designed IR illuminator, and a video decoder. The IR illuminator consists two sets of IR

LEDs, distributed evenly and symmetrically along the circumference of two coplanar concentric rings as shown in Figure 2.

Figure 2. An actual photo of the two rings IR illuminator configuration and the CCD camera

The video decoder we developed separates the input interlaced image into two fields, even and odd, and uses the signal to alternately turn the outer and inner IR rings of the illuminator on to produce the dark and bright pupil image on the even and odd field images respectively as shown in Fig. 3. The bright and dark pupil effects are subsequently exploited for accurate and robust eyes tracking in real time.



(a) (b)

Figure 3. The bright eye image (a) and the dark eye image (b), resulted by illuminating the face with LEDs in inner ring and outer ring respectively.

3.2 Eye Detection and Tracking

Eye activities can reveal antinational mechanisms and provide a window into one's cognitive and psychomotor capabilities. People experiencing fatigue or drowsiness tend to have abnormal eye activities such as slower eye blinks, longer eye closure duration, more eyelid droops, diminished eye blink frequency, and less pupil movement. Eye detection and tracking is therefore important to understand eye activities. Our research has led to a computer vision algorithm that can robustly and accurately detect and track eyes in real time and compute various parameters related to eyelid and pupil movement. By combining the latest technologies in appearance-based object recognition and tracking with active IR illumination, our eye tracker can robustly track eyes under variable and realistic lighting conditions and under various face orientations. In addition, our integrated eye

tracker is able to handle occlusion, glasses, and to simultaneously track multiple people with different distances and poses to the camera. Figure 4 summarizes our eyes detection and tracking algorithm. Details on our eye tracking algorithms can be found in [19].

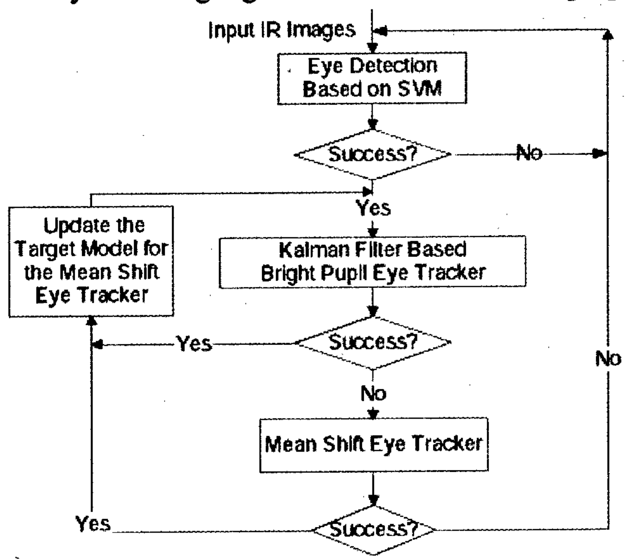


Figure 4. Eyes detection and tracking system flowchart

The primary goal of eye tracking is to monitor eyelid movement. Various parameters have been proposed to measure eyelid movement such as blink frequency, blink speed, eye closure duration, and PERCLOS. For this research, we focus on real-time computation of two eyelid movement parameters: PERCLOS and AECS. PERCLOS measures the percentage of eyelid closure over the pupil over time. A recent study by the Federal Highway Administration [4,17] shows, among many drowsiness-detection measures, PERCLOS was found to be the most reliable and valid ocular measure of a person's alertness level. AECS computes the average eye closure and opening speed, as determined by the amount of time needed to fully close/open the eyes. Our preliminary study indicates that the eye closure speed is distinctively different between a drowsy and alert subject. This may be explained by the tired muscle near the eyes for a person in fatigue. Figure 5 shows the detected eyes and the real time display of the running average measurements of PERCLOS and AECS over time.

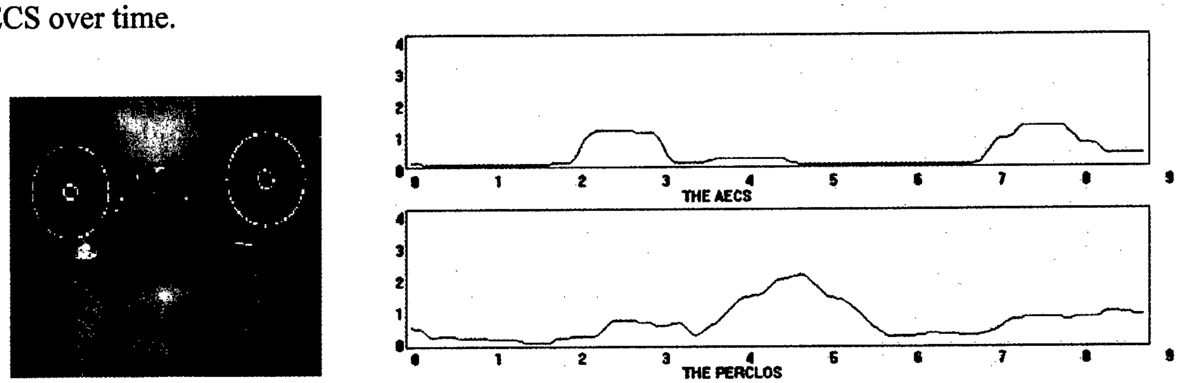


Figure 5. The detected eyes (right) and real time computing and displaying the running average measurements of PERCLOS and AECS.

3.3 Gaze Detection and Tracking

Gaze has the potential to indicate a person's level of vigilance. A fatigue individual tends to have a narrow and/or slow gaze movement. Gaze may also reveal one's needs and attention. Gaze estimation is important not only for fatigue detection but also for identifying a person's focus of attention, which can be used in the area of human-computer interaction.

The current remote gaze trackers work well only for a static head, which is a rather restrictive constraint on the part of the user. This poses a significant hurdle for practical application of the system. Another serious problem with the existing eye and gaze tracking systems is the need to perform a rather cumbersome calibration process for each individual. Often re-calibration is even needed for the same individual who already underwent the calibration procedure, whenever his/her head moves. In view of these limitations, we present a gaze estimation approach [32] that accounts for both the local gaze resulted from pupil movement and the global gaze resulted from the head movement. The global gaze and local gaze are combined together to obtain the precise gaze of the user. Our approach, therefore, allows natural head movement while still estimating gaze accurately. Another effort is to make the gaze estimation calibration free. New users or the existing users who have moved, do not need undergo a personal gaze calibration before using the gaze tracker. This is made possible by the use of Generalized Regression Neural Networks (GRNNs) to map pupil properties to screen coordinates. Therefore, the proposed gaze tracker can perform robustly and accurately without calibration and under natural head movements. A US patent is pending for our gaze detection and tracking algorithm. An overview of our gaze estimation algorithm is shown in Figure 6. More on our gaze estimation technique may be found in [32].

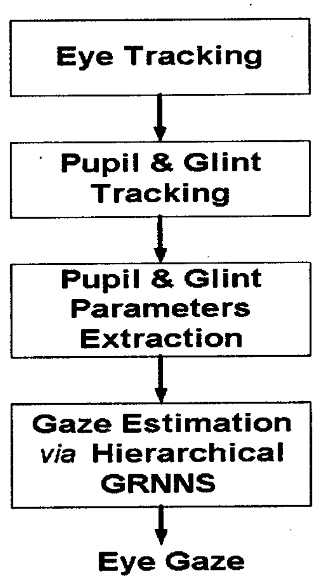


Figure 6 Major components of the proposed gaze estimation system,

Two gaze parameters are computed to characterize alertness. They are PerSac and GazeDis. While PerSac computes the percentage of saccade eye movement over time, GazeDis measures spatial fixation distribution over time. It is assumed that an alert person has a larger visual awareness and

experiences more frequent saccade movements than those of a sleepy person. Figure 7 gives a running average plot of PerSac.

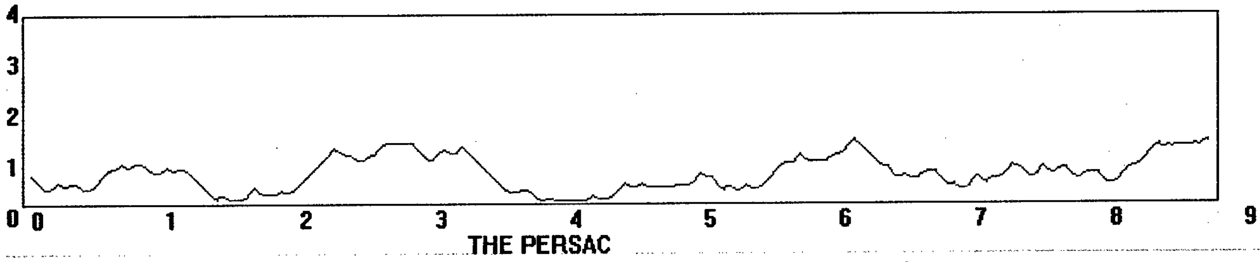


Figure 7 Plot of PERSAC parameter over time

3.4 Face Pose Tracking

Besides eye activities, head movement like nodding or inclination or frequent head tilts is a good indicator of a person's fatigue or the onset of fatigue [1]. In fact, irregular head movement (e.g., nodding) often occurs with people in fatigue. Head movement parameters such as head orientation, movement speed, frequency, etc. could potentially indicate one's level of vigilance.

Our research in this area focuses on developing computer vision algorithms for real time face detection and 3D face pose tracking from a monocular camera. So far, this research has led to the development of four different algorithms [8,11,12,37]. The first method focuses on determining face orientation by modeling face as an ellipse and determining face orientation based on the ellipse distortions. The second algorithm performs face orientation classification by performing a wavelet transform on the image and uses the wavelet coefficients (those sensitive to face orientation) to discriminate different face orientations. The third one performs face orientation determination based on the relationship between face orientations and the geometric properties of pupils. The fourth algorithm [37] assumes face can be modeled by a planar rectangle. Face detection and tracking is performed simultaneously. This method allows to real time estimate 3 face angles: pan, tilt, and swing. This algorithm is more robust and accurate and is finally adopted for face pose tracking. Figures 8 presents results of the face orientation estimation algorithm on an image sequences, with the estimated face normal as indicated by the white line near the nose.

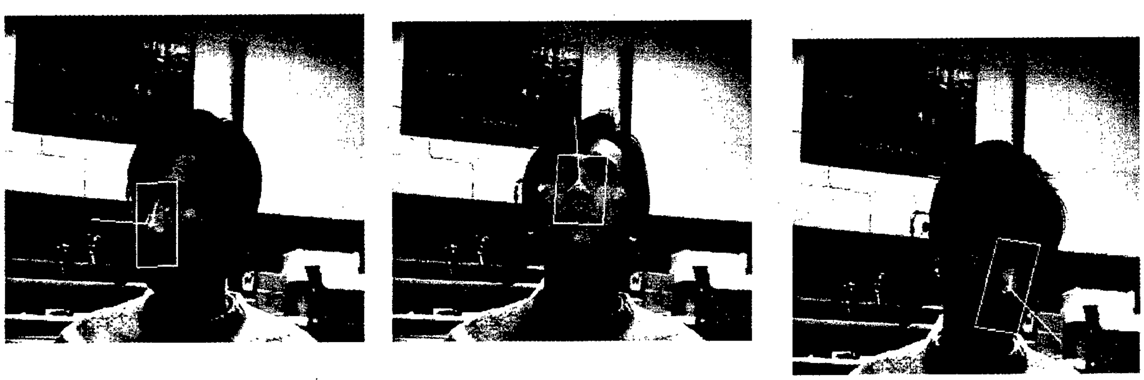


Figure 8: The estimated face normal in different frames under different face orientations. The white line vector represents the 3D face normal estimate

Figure 9 plots of the 3 face angles over time.

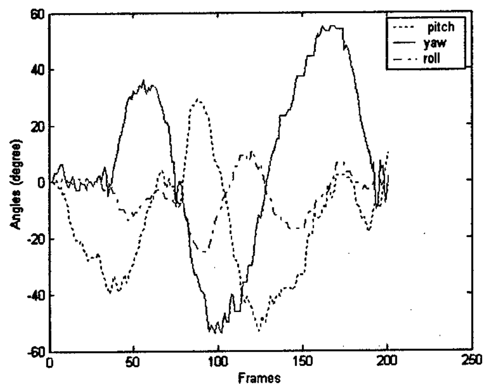


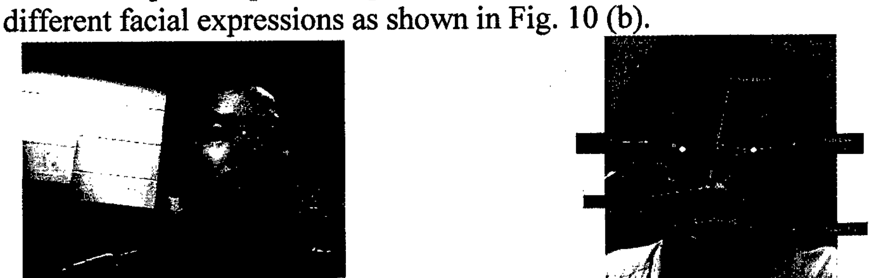
Figure 9: Face angles change over time

The parameter we compute to relate head movement to fatigue is PerNod, which computes the frequency of head tilt over time.

3.5 Facial Expressions Recognition

Facial expressions such as yawning or lagging muscles or expressionless are all visual symptoms of fatigue. The problem of analyzing facial expressions has become very important towards realizing a variety of applications such as advanced man-machine interfaces, human cognitive state monitoring, and visual communication systems. In general, people tend to exhibit different facial expressions under different levels of vigilance. For example, a drowsy person can be characterized by the slackness of the face muscles, the drooping of the upper eyelids, and frequent yawning. We believe that facial expressions provide yet another source of information to characterize a person's alertness.

We developed algorithms [34,38] for automatic facial feature tracking and facial expressions analysis. To characterize facial expression, our algorithm first identifies a few facial feature points (22) obtained by feature extraction in the frequency domain via Gabor filtering, guided by the bright pupils detected from eye tracking algorithm. The feature points are located near eyes, nose, and mouth as shown in Fig. 10 (a). The spatial semantics among the tracked features are then used to characterize facial expressions. The features are spatially related by graphs, with each feature point representing the node of the graph. The graph is elastic in that it deforms under different facial expressions as shown in Fig. 10 (b).



(a) (b)

Figure 10 (a) 22 facial features tracked; (b) the local graphs for facial expressions analysis

Different facial expressions can therefore be captured by different spatial configuration of the future points or the elastic graphs. Figure 11 represents two different facial expressions with detected feature points superimposed.

(a) (b)

Figure 11: A face with two different facial expressions: (a) serious, and (b) drowsy and yawning. Facial expression can be characterized by spatial configuration of the feature points, which are superimposed on the original images and are represented by elastic graphs.

Based on the detected and tracked facial features over time, we developed a new approach to facial expression understanding in image sequences [38]. We propose a stochastic framework, based on combining the Dynamic Bayesian Networks (DBNs) with Ekman's FACS coding [47], for expression representation and recognition. The DBNs has the expressive power to capture the dependencies, uncertainties and temporal behaviors exhibited by facial expressions, so that dynamic behaviors of facial expressions can be well modeled. The recognition of facial expressions is accomplished by fusing not only from the current visual observations, but also from the previous visual evidences. Consequently, the recognition becomes highly robust and accurate through the modeling of temporal behavior of facial expression. Figure 12 (top) shows an image sequence with two different facial expressions (happy and surprise) varying in intensity from frame to frame. Figure 12 (bottom) plots the probability of each of the two facial expressions over time.

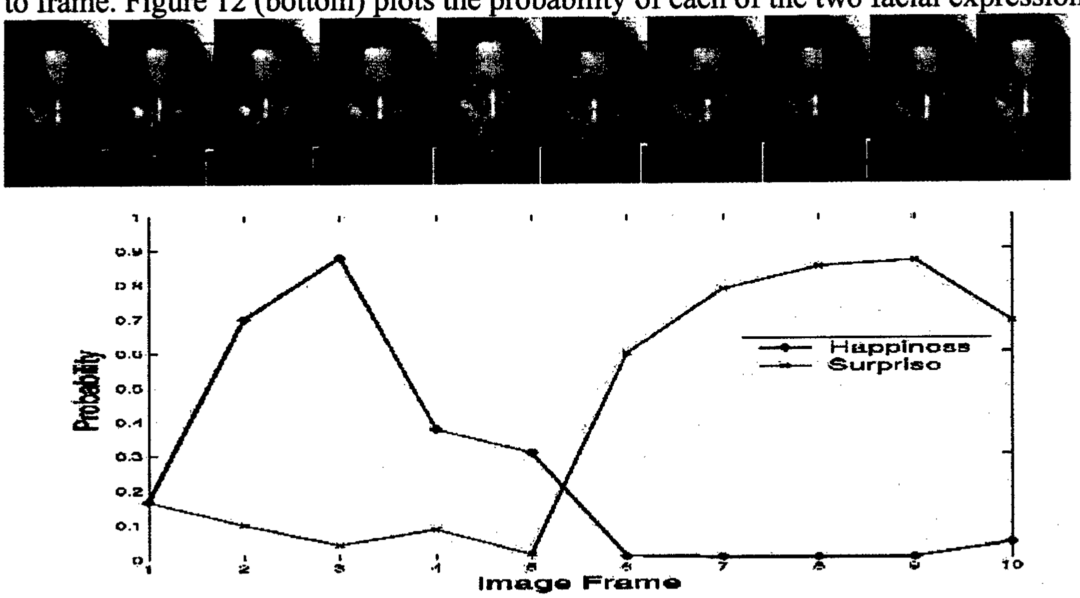


Figure 12 Facial expression recognition in a sequence. Top: two expressions (happy and surprise) vary in intensity. Bottom: the estimated probability for each expression.

For fatigue monitoring, we are particularly interested in detecting yawning. Yawning is characterized by mouth movement. Figure 13 plots mouth openness over time. A parameter PerYwan can be computed from the mouth openness to measure the frequency of yawning.

Figure 13 Plot of the openness of the mouth over time. The bumps are the detected yawns.

4. Fatigue Modeling via Bayesian Networks

4.1 Motivation and Introduction

The results of visual cues extraction are the extracted visual fatigue measures. These extracted fatigue measures are from different visual behaviors characterizing human fatigue from different perspectives. Furthermore, by the nature of process used in extracting information form the images, uncertainties exist concerning the properties of the extracted visual fatigue measures. The extracted fatigue measures in support or denial of a particular level of fatigue are therefore partial or incorrect or even conflictive with each other. On the other hand, all those visual cues, however imperfect and diverse they are, if combined, can provide an accurate fatigue characterization.

In addition to the extracted visual fatigue measures, there exist relevant contextual information that may lead to human fatigue. The specific prior contextual information such as physical fitness, sleep history, ambient temperature, and time of day are all important circumstantial factors, which, if known, will significantly improve the fatigue prediction accuracy. The use of different visual fatigue measures, the uncertainties associated with the extracted fatigue measures, and the incorporation of contextual information requires a mechanism to systematically integrate the diverse sources of evidences in a principled manner so that a consistent overall evaluation of a person's vigilance level can be achieved. By aggregating evidences from multiple sources into one representative format, we can reduce the uncertainty and resolve the ambiguity present in the information from a single source. The fusion process, thus, may solve the problem of local conflicting decisions and enhance the global accuracy for overall results. Information fusion and evidence integration are realized using the Bayesian probabilistic networks. A Bayesian network provides a mathematically coherent and a sound basis for uncertainty representation and for aggregating evidences. The goal of Bayesian inference is to identify a person's fatigue level that could best explain the observed visual behaviors and the available contextual information.

4.2 Fatigue Modeling Using Bayesian Networks

A Bayesian Network (BN) provides a mechanism for graphical representation of uncertain knowledge and for inferring high level activities from the observed data. Specifically, a BN consists of nodes and arcs connected together forming a directed acyclic graph (DAG) [20]. Each node can be viewed as a domain variable that can take a set of discrete values or a continuous value. An arc represents a probabilistic dependency between the parent node and the child node. Since BN was developed to model the distribution processing in reading comprehension in the late

of 1970's, numerous studies have been conducted and many systems have been constructed based on this paradigm in a variety of application areas, including industrial applications, military, medical diagnosis and commercial applications [21,22,23].

The main purpose of a BN model is to infer the unobserved events from the observed and contextual data. So, the first step in BN modeling is to identify those hypothesis events we want to infer and group them into a set of mutually exclusive events to form the target hypothesis variable. The second step is to identify the observable data that may reveal something about the hypothesis variable and then group them into information variables. There are also other hidden states which are needed to link the high level hypothesis node with the low level information nodes. For fatigue modeling, fatigue is obviously the target hypothesis variable that we intend to infer while other contextual factors, which could cause fatigue, and visual cues, which are symptoms of fatigue, are information variables. Of many factors that can cause fatigue, the most significant ones are sleep history, circadian, work condition, work environment, and physical condition. These contextual factors can be further broken down as follows. The most profound factors that characterize work environment are temperature, weather and noise; the most significant factors that characterize physical condition are age, sleep disorders and food; the factors affecting work conditions include workload and type of work. Furthermore, factors affecting sleep quality include sleep environment and sleep time. The sleep environment includes random noise, background light, heat and humidity around the bed. From the computer vision module, we can obtain several visual fatigue parameters to characterize eyelid movement (PERCLOS and AECS), gaze (PerSac and GazeDis), head movement (PerNod), and the facial expression (PerYwan). Putting all these factors together, the BN model for fatigue is constructed as shown in Fig. 14. The target node is fatigue and the nodes above the target node represent various major factors that could lead to one's fatigue. They are collectively referred to as the contextual information. The nodes below the target node represent visual observations from the output of our computer vision system. These nodes are collectively referred to as the observation nodes.

4.3 Construction of conditional probability table (CPT)

Before using the BN for fatigue inference, the network needs be parameterized. This requires learning the prior probability for the root nodes and the conditional probabilities for the links from the training data. For this research, training data were obtained from three different sources including data obtained from our human subject study, data from several large-scale subjective surveys [26,27,28,29], and some subjective numbers from experts.

From our human subjects study, we collected a large amount data from 16 experiments for 8 subjects. Data consists of TOVA task performance data and visual parameters computed by our computer vision system. TOVA performance lapses can be used as a ground-truth measure of alertness while the visual observations can serve as the sensory observations. These data are used to train the lower part of the fatigue model. The upper part of model is parameterized based on the data from the surveys, despite their subjectivity. Since these surveys were not designed for the parameterization of our BN model, not all needed probabilities are available and some conditional probabilities are therefore inferred from the available data using the so-called *noisy-or* principle [24]. Still some prior or conditional probabilities are lacking in our model, they are obtained by subjective estimates methods [24]. With this, all the prior and conditional probabilities in our BN model are obtained. Details on the learning of CPTs and the final numbers may be found in [41].

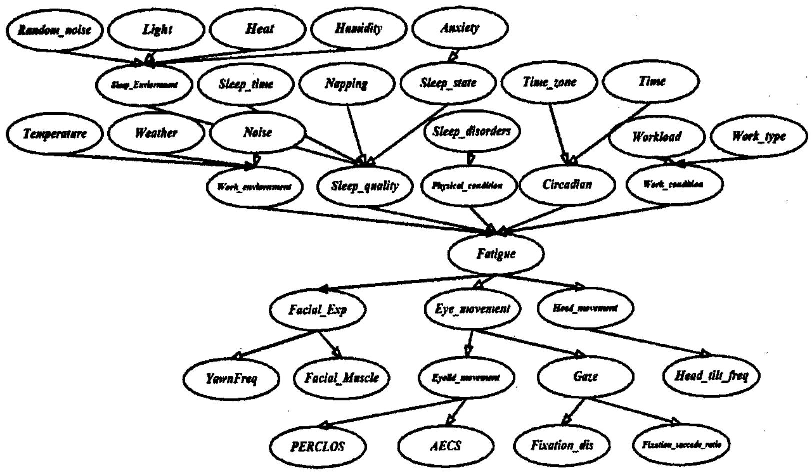


Fig. 14. A Bayesian network model of human fatigue

4. 4 The Experimental Results

Given the parameterized model, fatigue inference can then commence upon the arrival of visual evidences via belief propagation. MSBNX software [25] is used to perform the inference and both top-down and bottom-up belief propagations are performed. Since it is impossible to enumerate all possible input combinations, here we only simulate some typical combination of evidences and the results are summarized in Table 1

Table 1: Fatigue Inference Results from the Bayesian Fatigue Model

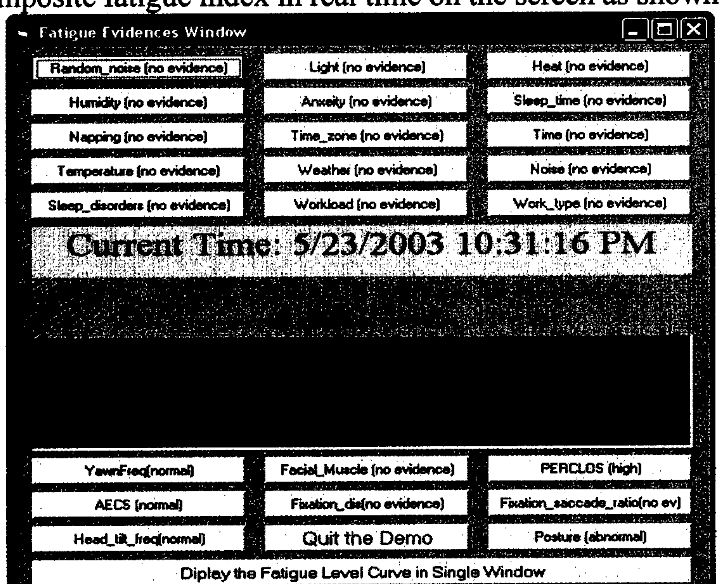
No.	Evidences Used	Posterior probabilities of Fatigue ('yes' state)
1	No any evidence	0.58
2	PerYawn (high)	0.82
3	PERCLOS (high)	0.89
4	PerSac (low)	0.84
5	PerNod (high)	0.73
6	Temperature (high)	0.72
7	Weather (abnormal)	0.73
8	Noise (high)	0.71
9	Age (>45 year)	0.60
10	Circadian (drowsiness)	0.77
11	Sleep disorder (yes)	0.60
12	Food (hungry)	0.63
13	Workload (heavy)	0.72
14	Type work (tedious)	0.74
15	Worry (yes)	0.69
16	Random Noise (often)	0.60
17	Light (yes)	0.60
18	Heat (high)	0.60
19	Sleep time (loss)	0.68
20	PERCLOS (high), PerSac (low)	0.95
21	PERCLOS (high), PerYawn (high)	0.97
22	PerSac (low), PerYawn (high)	0.94
23	PerSac (low), PerNod (high)	0.91
24	PerNod(high), PerYawn (high), PerSac (low)	0.95
25	PerSac (low), Circadian (drowsiness)	0.92
26	PerYawn (high) , Food (hungry), Random Noise (yes), Temperature (high), Type work (tedious)	0.95
27	PERCLOS (high), Random Noise (often), Temperature (high), Worry (yes)	0.96
28	PerNod (high), PerYawn (high), Random Noise (often), Temperature (high), Worry (yes)	0.96

29	PerNod (high), PerSac (low), Random Noise (often), Sleep disorder (yes), Temperature (high)	0.96
30	Age (>45 year), Circadian (drowsiness), Food (hungry), Heat (high), sleep humidity (high), Sleep disorder (yes), sleep time (loss), Type work (tedious), Weather (abnormal), Workload (heavy), Worry (yes)	0.96

From Table 1, we can see that the prior probability of fatigue (e.g. when there is not any evidences) is 0.58 (Row. #1). The observation of single visual evidence (Rows #2-#5) does not provide conclusive finding since the estimated fatigue probability is less then the critical value 0.95 (arbitrarily chosen), even the observation of high PERCLOS measurement (Row #3) can not produce sufficient confidence in fatigue. Similarly, the presence of a single contextual factor (Row #6-#19) cannot produce high probability of fatigue. On the other hand, the combination of two visual evidences (Row. #20~#23),) leads to a fatigue probability close to or higher than 0.95. Any combination of three visual cues guarantees the estimated fatigue probability exceeds the critical value (Row #24). The same can be achieved by combining visual evidences with contextual evidences (Row #26-#29). This demonstrates the importance of contextual information. In fact, the simultaneous presence of all contextual evidences only almost guarantees the occurrence of fatigue (Row #30). These inference results, thought preliminary and synthetic, demonstrate the utility of the proposed framework for predicting and modeling fatigue by pooling information from different sources.

4.5 System Integration

The vision module and fatigue model is subsequently integrated to produce the prototype fatigue monitor. For this, an interface program has been developed to connect the output of the computer vision system with that of the information fusion engine. Upon the arrival of new evidences from the vision module, the interface program instantiates the evidences of the fatigue network, which then performs fatigue inference. The interface program then displays and plots the composite fatigue index in real time on the screen as shown in Figure



15.

Figure 15 The display panel of the interface program that integrates the vision module with the information fusion module. The program displays and plots the fatigue score (the curve in the middle) in real time.

5. System Validation

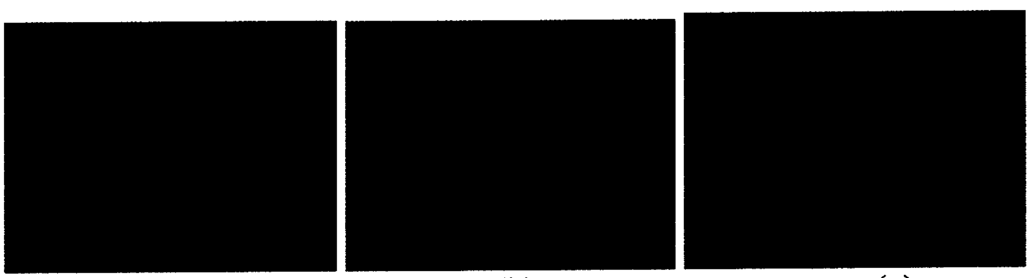
The last part of this research is to experimentally and scientifically demonstrate the validity of the computed fatigue parameters as well as the composite fatigue index.

5.1 Validation of the measurement accuracy

Here, we present results to quantitatively characterize the measurement accuracies of our computer vision techniques in measuring eyelid movement, gaze, face pose, and facial expressions. The measurements from our system are compared with those obtained either manually or using conventional instruments.

5.1.1 Eye detection and tracking accuracy

This section summarizes the eye detection and tracking accuracy of our eye tracker. For this study, we randomly selected an image sequence that contains 13,620 frames, and manually identified the eyes in each frame. The manually labeled data serves as the ground-truth data and are compared with the eye detection results from our eye tracker. The study shows our eye tracker is quite accurate, with a false alarm rate of 0.05% and a misdetection rate of 4.2%. Further study shows that the misdetections can be broken down into three cases. In case 1, the eye is fully open, but our tracker fails to detect the eyes. This accounts for about less than 1% of misdetections. In cases 2 and 3, eyes are misdetected in the frames just prior to or after the eye is closed as shown below.

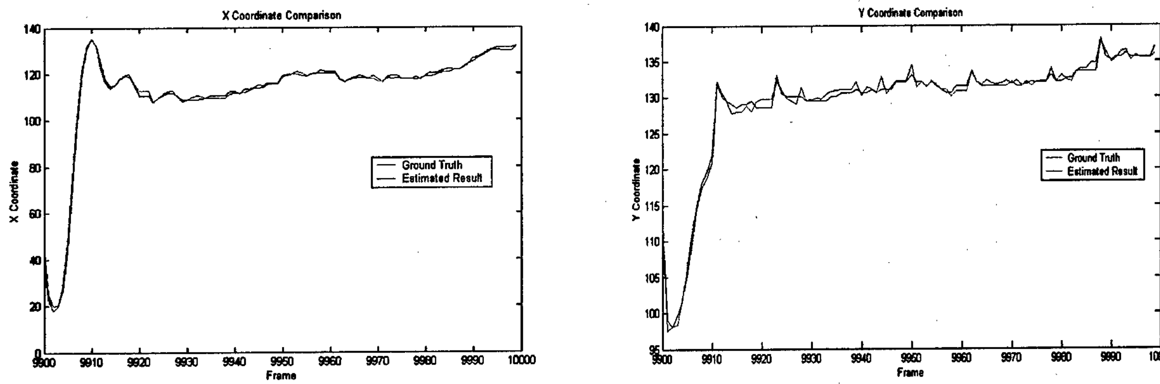


(a) (b) (c)

Figures 16: Cases of eye misdetections: (a) eye is fully open; (b) eye begins to close, (c) eye begins to open after closure.

5.1.2 Eye detection and eye parameter estimation accuracy

In this experiment, we studied the positional accuracy of the detected eyes as well as the accuracy of the estimated pupil size. The ground-truth data are produced by manually determining the locations of the eyes in each frame as well as the size of the pupil. The size of the pupil is determined by manually selecting a few points along the boundary of the pupil and then performing an ellipse fitting on the selected points. The pupil size is then characterized by the ratio of major axis length to that of minor axis. The ratio is also used to characterize the degree of eye opening. Figures 17 and 18 summarize the comparison results. It is clear from Figure 17 that the detected eye positions match very well with manually detected eye positions, with a RMS position errors of 1.09 and 0.68 pixels for x and y coordinates respectively.



(a) (b)

Figure 17 The estimated eye positions versus the manually detected eye positions for 100 random selected consecutive frames: (a) x coordinates and (b) y coordinates.

Figure 18 shows the estimated pupil size (ratio) versus the manually determined the pupil size for the same image sequence. The two curves basically match, with a RMS error of 0.0812. The discrepancies are primarily caused by the different methods used to estimate the pupil ratio. The automated method computes the pupil ratio based on all pixels of the pupils while the manual method uses only the boundary pixels. In addition, inaccuracy and inconsistency in selecting boundary points by the manual method further contributes to the differences.

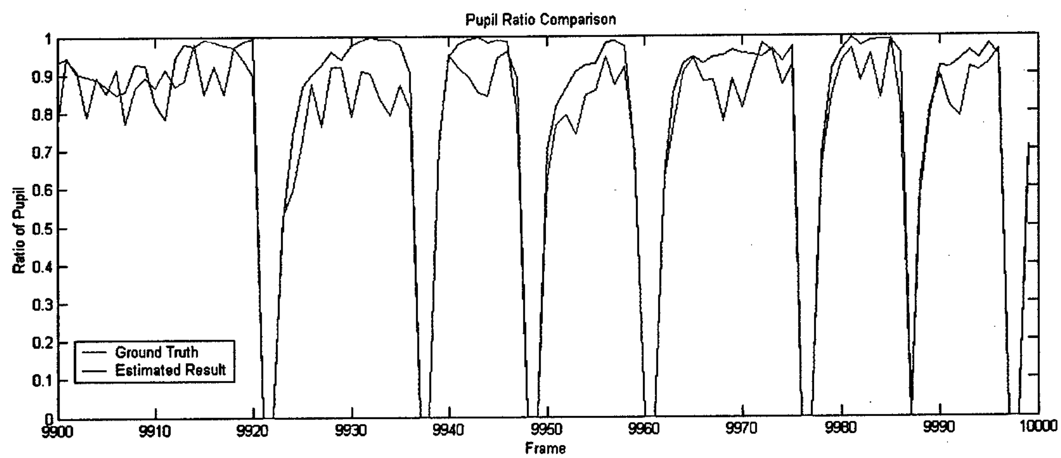
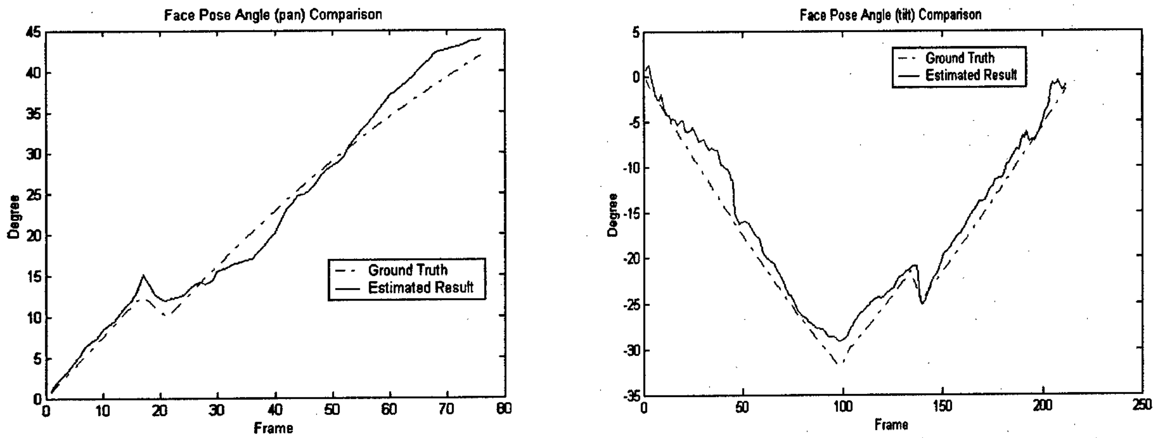


Figure 18: The estimated pupil size versus the manually determined pupil size for 100 random frames.

5.1.3 Face pose parameters accuracy

Here, we present the experimental results that validate the accuracy of our face pose estimation. Our face pose estimation computes in real time 3 face angles, pan, tilt, and swing. To study their accuracy, we use a head-mount head tracker that tracks head movements. The output of the head-mount head tracker is used as the ground-truth. Figure 19 visually plots the tracking results of our face tracker versus that of the head tracker. It is apparent that qualitatively, the two trackers match each other pretty well.



(a) (b)

Figure 19 The estimated pan angle (a) and tilt angle (b) versus the angles computed by the head tracker for 80 frames

Quantitatively, the RMS errors for the pan and tilt angles are 1.92 degrees and 1.97 degrees respectively. This experiment demonstrates that our face pose estimation technique is sufficiently accurate.

5.2 Validation of fatigue parameters and the composite fatigue score

To study the validity of the proposed fatigue parameters and that of the composite fatigue index, we performed a human subject study. The study included a total of 8 subjects. All are healthy including two females. The oldest subject is 65 while the youngest subject is 21 year's old. Two test bouts were performed for each subject. The first test was done when they first arrived in the lab at 9 pm and when they were fully alert. The second test was performed about 12 hours later early in morning about 7 am the following day, after the subjects have been deprived of sleep for a total of 25 hours.

During the study, the subjects are asked to perform a TOVA test. The TOVA test consists of a 20-minute psychomotor test, that requires the subject to sustain attention and respond to a randomly appearing light on a computer screen by pressing a button. TOVA test was selected as the validation criterion because flying or driving is primarily a vigilance task requiring psychomotor reactions, and psychomotor vigilance. Various performance measures are used to evaluate the subject's performance in 2 seconds interval including response time, omission and commission errors. For each subject, we collect the following data: visual data (eyelid movement, gaze, facial expressions, and face pose), TOVA task performance measures, and EEG.

5.2.1 TOVA Performance Lapses Versus Fatigue

TOVA performance lapses occur when the subject's response time to the target signal is over 500 ms or when the subject fails to responds to the signal (omission). In this experiment, we study the average TOVA performance lapses over all the subjects for the two different bouts. The average TOVA performance lapses for bout 1 is 26 times while the average lapses for bout 2 is 56 times, apparently a significant increase in the number of performance lapses for the morning bout. In addition, the response time also varies between the two bouts. Figure 20 plots the response time for the two bouts for two subjects. It shows the response time is generally longer for the morning

bout for both subjects. This demonstrates that TOVA performance response time correlates well with level of fatigue.

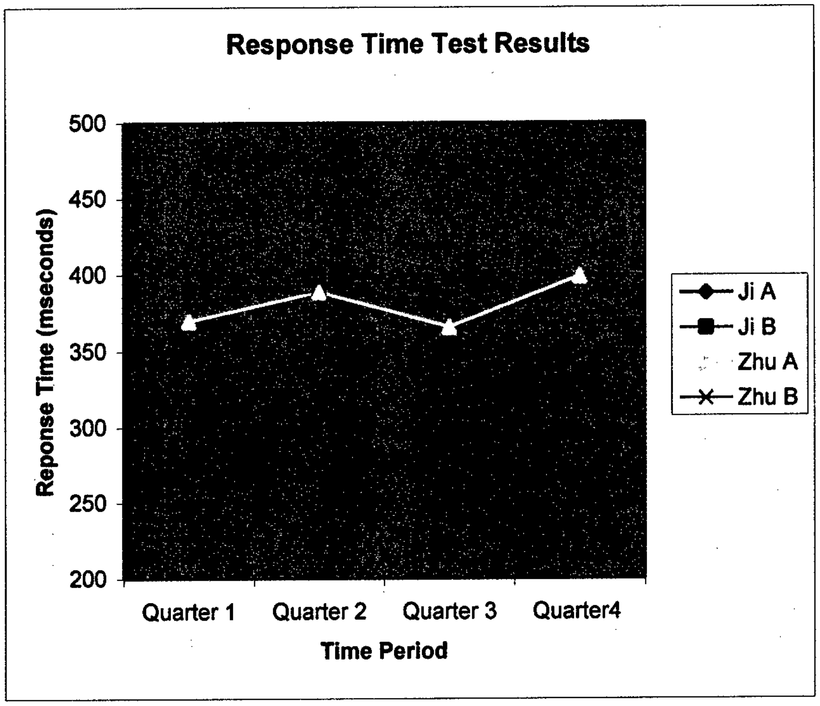


Figure 20 Plots of TOVA response time for two subjects (Ji and Zhu) for two study bouts: one in the evening (A) when the subject is awake and the other is in the early morning (B) when the subject has been deprived 16 hours of sleep.

5.2.2 Validation of the PERCLOS measure

Here we present results to show the correlation of the computed PERCLOS with the TOVA performance lapses and with level of fatigue. Figure 21 plots TOVA performance lapses v.s. PERCLOS measurements. It is clear that most of the performance lapses happen near the peaks of the PERCLOS. This demonstrates the correlation between the performance lapses and high PERCLOS measurements.

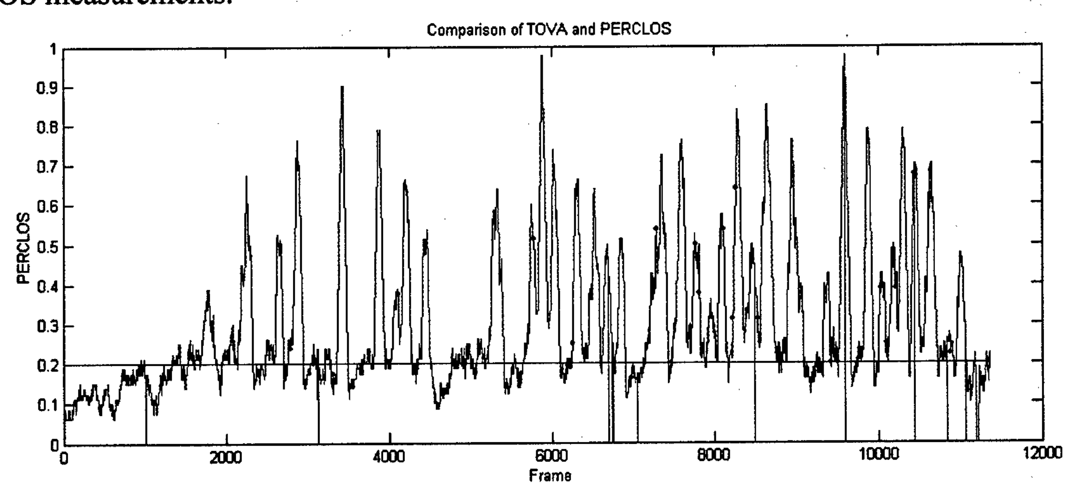


Figure 21a: TOVA performance lapses (blue dots) superimposed on PERCLOS plot for the entire 20 minutes for a subject.

If the PERCLOS threshold is set at 0.2, then the agreement rates for the figures above is 0.76, i.e., 76% of the performance lapses occur near the peaks of PERCLOS.

To further study the correlation between PERCLOS and the reaction time, we plotted the average reaction times versus average PERCLOS measurements as shown in Figure 21b. The figure clearly shows the approximate linear correlation between PERCLOS and the TOVA response time. This experiment once again demonstrates the validity of PERCLOS in quantifying vigilance, as characterized by TOVA response time.

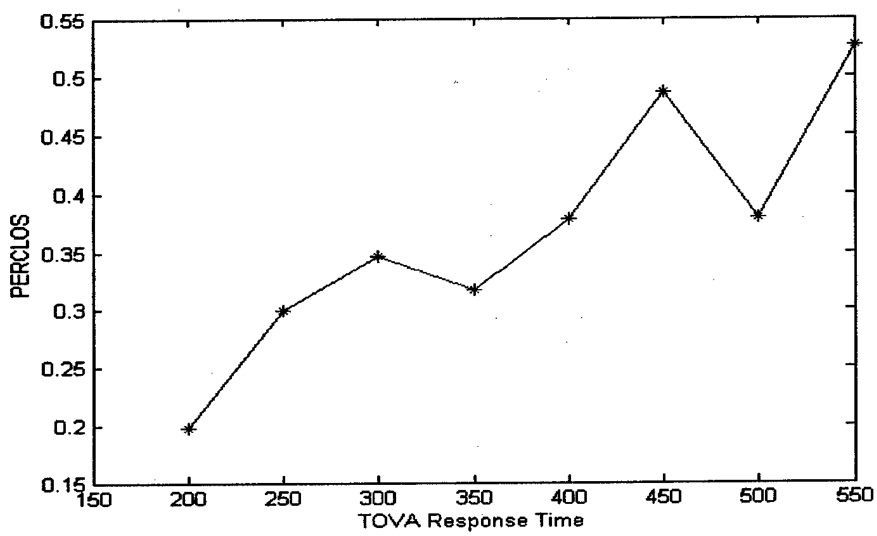


Figure 21b: PERCLOS versus TOVA response time. The two parameters are clearly correlated almost linearly. A larger PERCLOS measurement corresponds to a longer reaction time.

Finally, we want to demonstrate the correlation between PERCLOS and fatigue. For this, we compared the PERCLOS measurements for two bouts for the same individual. The comparison is shown in Figure 21c, where it is clear that the PERCLOS measurements for the night bout (when the subject is alert) is significantly lower than the morning bout (subject is fatigue). This not only proves the validity of PERCLOS to characterize fatigue but also proves the accuracy of our system in measuring PERCLOS.

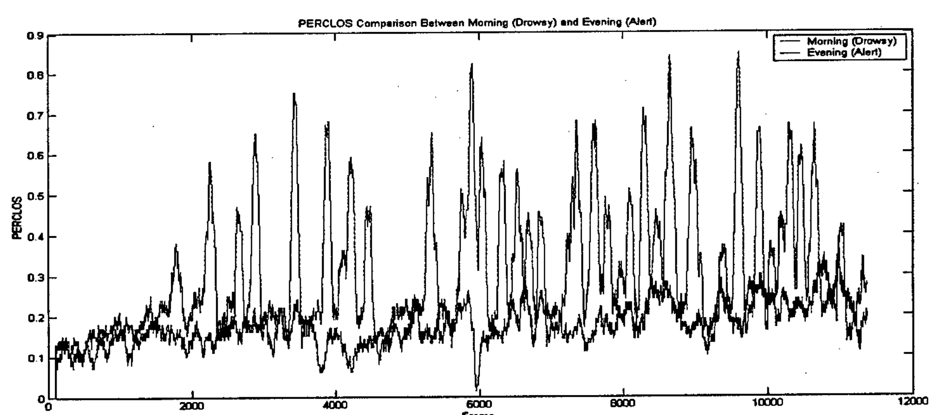


Figure 21c: PERCLOS measurements for evening (blue) and morning (red) bouts

5.2.3 Validation of AECS parameter

ACES represents the average eye closure and opening speed. In this experiment, we want to verify its validity as an ocular measure of human fatigue. Again, we plot AECS over the entire period and superimpose the TOVA performance lapses on the curve to see if they coincide with high values of AECS as shown in Figure 22.

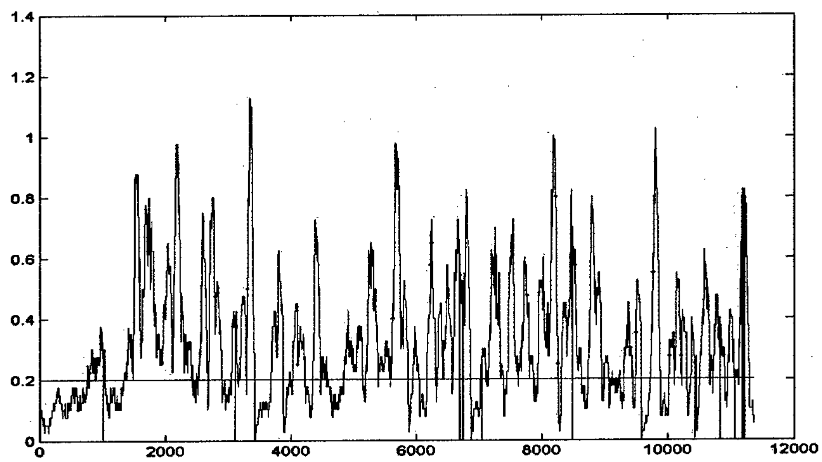


Figure 22a: TOVA performance lapses (blue dots) superimposed on AECS plot for the entire 20 minutes. It is clear that most of the performance lapses happen near the peaks of the AECS (corresponding to longer closure time). This demonstrates the correlation between the performance lapses and the high AECS measurements.

To further study the correlation between AECS and the TOVA response time, we plotted the average reaction times versus average AECS measurements as shown in Figure 22b. The figure clearly shows the approximate linear correlation between AECS and the reaction time. This experiment once again demonstrates the validity of AECS in quantifying vigilance, as characterized by TOVA response time.

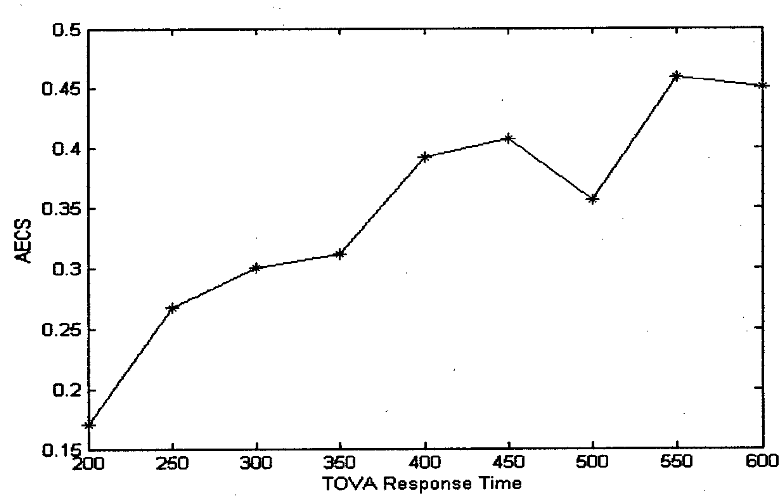


Figure 22b: AECS versus TOVA response time. The two parameters are clearly correlated almost linearly. A larger AECS measurement corresponds to a longer reaction time.

Finally, we want to demonstrate the correlation between AECS and fatigue. For this, we compared the AECS measurements for two bouts for the same individual. The comparison is shown in Figure 22c, where it is clear that the AECS measurements for the night bout (when the subject is alert) is significantly lower (faster) than the morning bout (subject is fatigue). This not only proves the validity of AECS to characterize fatigue but also proves the accuracy of our system in measuring AECS.

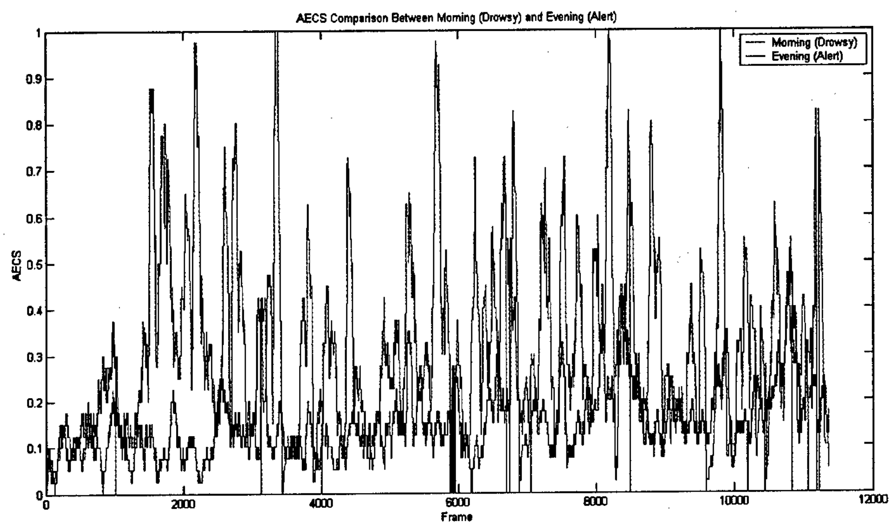


Figure 22c: AECS measurements for evening (blue) and morning (red) bouts

5.2.4 Validation of gaze parameter PerSac

PerSac represents the average saccade eye movement over time. In this experiment, we want to verify its validity as an ocular measure of human fatigue. Again, we plot in Figure 23a PerSac measure over the entire period and superimpose the TOVA performance lapses on the curve to see if they coincide with low values of PerSac. From Figure 23a, we can see that TOVA performance lapses mostly occur with low PerSac values, i.e., less saccade movement correlates with longer response time or slower reaction time.

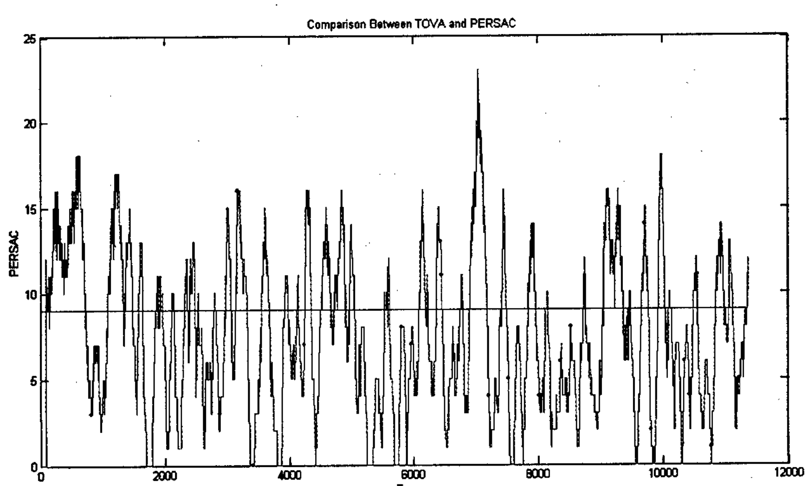


Figure 23a TOVA performance lapses (blue dots) superimposed on PerSac plot for the entire 20 minutes. It is clear that most of the performance lapses happen when PerSac measure is

low (corresponding to less saccade movement). This demonstrates the correlation between the performance lapses and low PerSac measurements.

To further study the correlation between PerSac and the TOVA response time, we plotted the average reaction times versus average PerSac measurements as shown in Figure 23b. The figure clearly shows the approximate negative linear correlation between PerSac and the response time. This experiment once again demonstrates the validity of PerSac in quantifying vigilance, as characterized by TOVA response time.

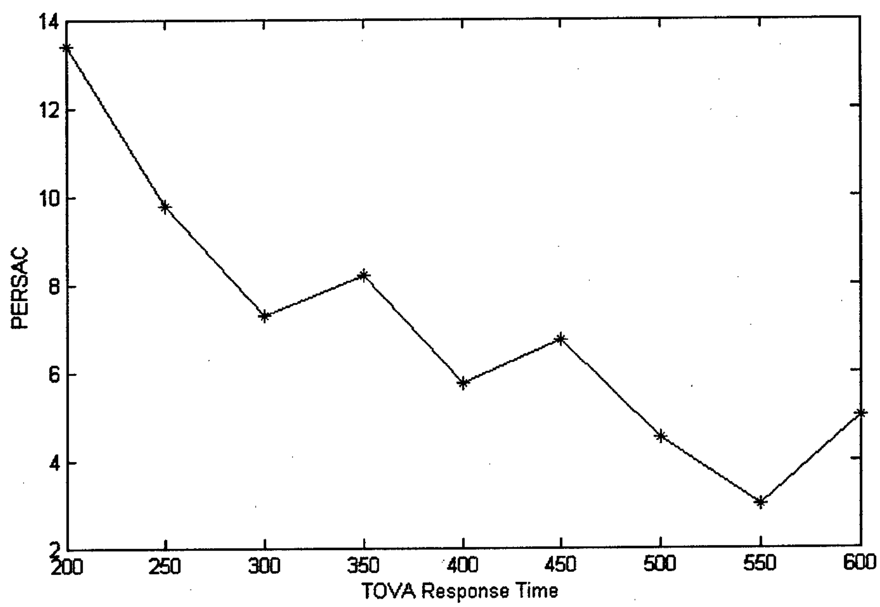


Figure 23b: PerSac versus TOVA response time. The two parameters are clearly correlated almost linearly. A smaller PerSac measurement corresponds to a longer response time.

Finally, we want to demonstrate the correlation between PerSac and fatigue. For this, we compared the PerSac measurements for two bouts for the same individual. The comparison is shown in Figure 23c, where it is clear that the PerSac measurements for the night bout (when the subject is alert) is significantly larger (more saccade movements) than the morning bout (subject is fatigue). This not only proves the validity of PerSac to characterize fatigue but also proves the accuracy of our system in measuring PerSac.

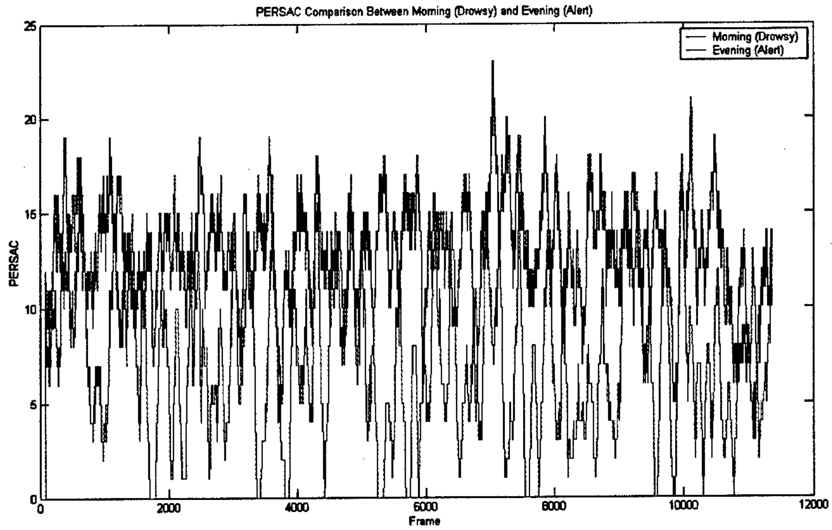


Figure 23c: PerSac measurements for evening (blue) and morning (red) bouts

5.2.5 Validation of the composite fatigue index

Here we studied the response time versus the composite fatigue index computed by our fatigue monitor. The results are plotted in Figure 24, which clearly shows that the composite fatigue score (based on combining different fatigue parameters) highly correlates with the subject's response time.

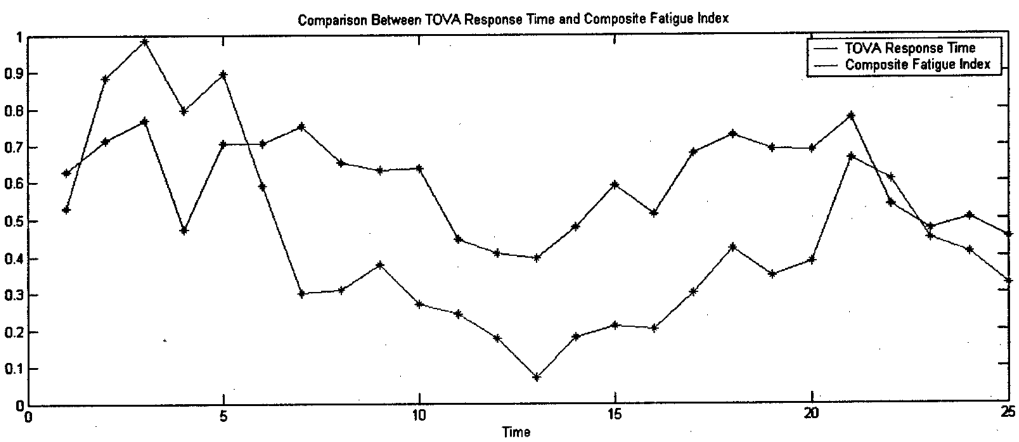


Figure 24: The estimated composite fatigue index (blue) versus the normalized TOVA response time. The two curves track each other well.

It is clear that the two curves' fluctuations match well, proving their correlation and co-variation. In the figures below, we try to demonstrate the co-variation and correlation between the composite fatigue index and the 3 individual fatigue parameters: PERCLOS, ACES, and PERSAC.

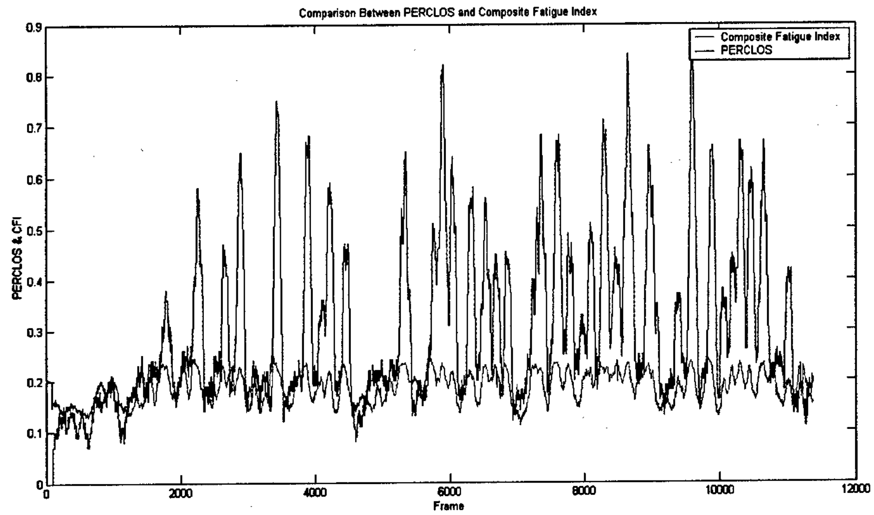


Figure 25 Perclos (blue) versus the composite fatigue score (red). They apparently track each other well.

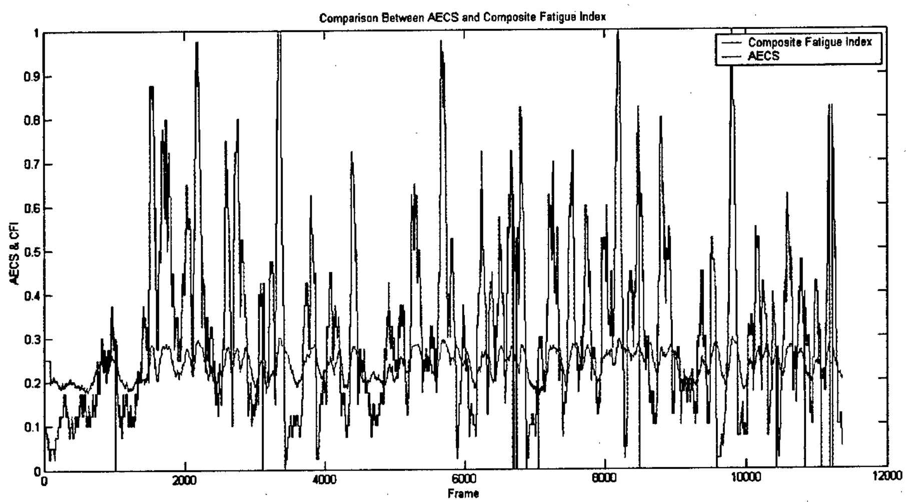


Figure 26 AECS (blue) versus the composite fatigue score (red). They apparently track each other well.

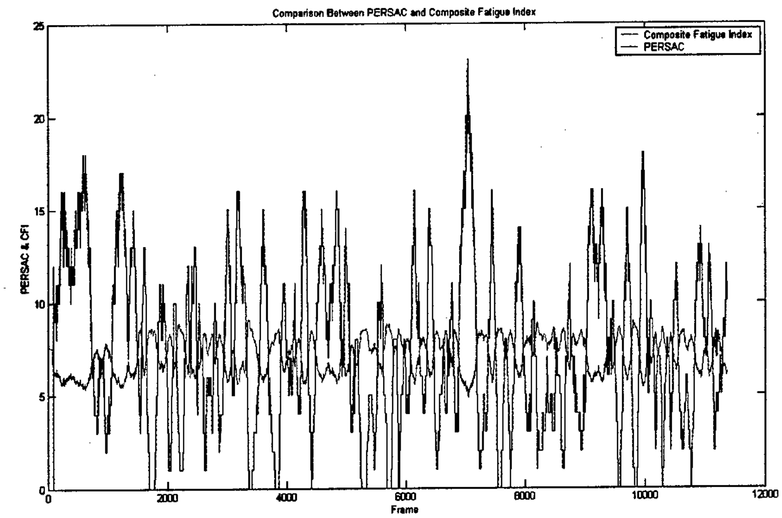


Figure 27 PerSac (blue) versus the composite fatigue score (red). They apparently negatively track each other well.

6. Transitions, invention, and media coverage

Through this research, we have been able to generate additional funding for this and related research. Specifically, we receive funding from Honda, Darpa, and ONR. Honda has been supporting this project for more than 2 years. Negotiation is currently under way to continue supporting this research in Phase 3. Recently, we have proposed to extend the fatigue monitoring to human emotion recognition. This effort is currently being funded by ONR/Darpar for 1.3 million dollars for 4 years. This research has so far yielded 14 publications and one patent (pending) including 4 journal publications. In addition, through the support of this project, two MS students in computer science have graduated. The project also supported a post-doctoral researcher.

Our research has been covered by various media outles including local newspapers, TV, and the New York Times. Below is a photo appearing in the business section of Aug. 26, 2003 issue of the New York Times.

Dr. Qiang Ji of Rensselaer Polytechnic Institute in Troy, N.Y., demonstrates a driver fatigue monitor.
We have also built a website to disseminate our research. The website includes published papers, video demos, and Internet resources related to eye tracking and human fatigue monitoring. The website URL is http://www.ecse.rpi.edu/~qji/Fatigue/fatigue.html

7. Conclusion

In this report, we summarize our efforts in developing real time non-intrusive technology for monitoring human fatigue. Through this research, we have developed state of the art technologies and a prototype fatigue monitor for real time non-intrusive human fatigue monitoring. Our contributions include: 1) the development of various computer vision techniques for real-time and non-intrusive extraction of multiple fatigue parameters related to eyelid movements, gaze, head movement, and facial expressions, 2) the development of a probabilistic framework based on the Bayesian networks to model and integrate contextual and visual cues information for accuracate and robust fatigue detection, and 3) systematic and scientific validation of the fatigue monitor. Experimental validation of our techniques using human subjects demonstrates the good measurement accuracy of our techniques. In addition, the study also verifies the validity of the proposed fatigue parameters as well as that of the composite fatigue index.

Our experience concluded that in order to monitor and predict human fatigue, the following must be satisfied. First, the technology must be non-intrusive. A technology, even with minimum intrusion, will have significant difficulty of acceptance in real world. Second, it is important to simultaneously extract multiple parameters and systematically combine them in order to obtain a robust and consistent fatigue characterization. Third, a fatigue model must be built to represent related knowledge and information and to infer a person's cognitive states from the observed sensory data. Our research basically covers all the three aspects. But significant research is ahead of us to further realize them. Future research includes 1) further development and improvement of the vision algorithms, 2) miniaturization of the hardware components of the fatigue monitor, 3) optimization of the software implementation, and 4) validation of our fatigue monitor with a field test.

References

[1] Anon. Proximity Array Sensing System: head position monitor/metric. *Advanced safety concepts, Inc. Sante Fe, NM87504.*

[2] Anon, 1999, Conference on ocular measures of driver alertness, Washington DC. April.

[3] S. Boverie , J. M. Leqellec and A. Hirl, 1998Intelligent systems for video monitoring of vehicle cockpit. *1998 International congress and exposition ITS: Advanced controls and vehicle navigation systems,* pages 1-5.

[4] David F. Dinges ,M. Mallis ,G. Maislin and J. W. Powell, 1998, Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management, *Department of Transportation Highway Safety publication 808 762*, April

[5] M. Kaneda et al, 1994, Development of a drowsiness warning system. *11th international conference on enhanced safety of vehicle, Munuch.*

[6] J. Feraric , M. Kopf and R. Onken, 1992, Statistical versus neural bet approach for driver behavior description and adaptive warning, *11th European annual manual*, pages 429-436.

[7] T. Ishii , M. Hirose and H. Iwata, 1987, Automatic recognition of driver's facial expression by image analysis, *Journal of JSAE*, 41(12):PAGES 1398-1403

[8] Qiang Ji and Rong Hu, 2002, 3D Face Pose Estimation and Tracking from a Monocular Camera, *Journal of Image and Vision Computing*, 20(7).

[9] Qiang Ji and Xiaojie Yang, 2001, Real time eye, gaze, and face pose tracking for monitoring driver vigilance, *in ICVS 2001: Second International Workshop on Computer Vision Systems,* Vancouver, Canada.

[10] Qiang Ji and Xiaojie Yang, 2002, Real-time eye, gaze, and face pose tracking for monitoring driver vigilance, Real-Time Imaging, Volume 8 , Issue 5, Pages: 357 - 377.
[11] Qiang Ji and Xiaojie Yang, 2002, Real Time 3D Pose Discrimination Based on Active IR Illumination,*16th International Conference on Pattern Recognition.*
[12] Mukesh Motwani and Qiang Ji, 2001, 3D Face Pose Discrimination Using Wavelets, *IEEE International Conference on Image Processing.*
[13] Reiner Onken, 1994, DAISY, an adaptive knowledge-based driver monitoring and warning system, *Proceedings of 1994 Vehicle navigation and information systems conference, Yokohama, Japan, Aug. 1994*, PAGES 3-10.
[14] H. Saito, 1994, T. Ishiwaka and M. Sakata and S. Okabayashi, Applications of driver's line of sight to automobiles-what can driver's eye tell, *Proceedings of 1994 Vehicle navigation and information systems conference, Yokohama, Japan, Aug. 1994*, PAGES 21-26.
[15] S. Saito, 1992, Does fatigue exist in a quantitative of eye movement ? *Ergonomics*, 35.
[16] H. Ueno, 1994, M. Kaneda and M. Tsukino, Development of drowsiness detection system, *Proceedings of 1994 Vehicle navigation and information systems conference, Yokohama, Japan, Aug. 1994*, PAGES 15-20.
[17] W. W. Wierville, 1994, L. A. Ellsworth and et al, Research on vehicle based driver status/performance monitoring: development, validation, and refinement of algorithms for detection of driver drowsiness, *National Highway Traffic Safety Administration Final Report: DOT HS 808 247.*
[18] K. Yammamoto and S. Higuchi, 1992, Development of a drowsiness warning system, *Journal of SAE of Japan*, 46(9).
[19] Zhiwei Zhu , 2002, Kikuo Fujimura and Qiang Ji, Robust Real Time Eye Tracking Under Various Light Conditions and Face Orientations, *Eye Tracking Research and Applications Symposium, New Orleans, LA, March.*
[20] Jordan, M. I. 1999, editor, Learning in Graphical Models, MIT press.
[21].H. Li, 1998, An Introduction to Belief Networks, Technical Research Report, CSHCN T. R. 99-31 (ISR T. R. 99-58), `http://www.isr.umd.edu/CSHCN/`.
[22] Sudeep Sarkar and Kim L. Boyer, 1993, Integration, inference, and management of spatial information using Bayesian networks: Perceptual organization, IEEE Transactions on Pattern Analysis and Machine Intelligence, p256-274, vol. 15, No. 3.
[23] R. Fung,1988, Structure composing for situation assessment, Ph.D thesis, Engineering Economic Systems, Stanford University.
[24] Jensen, F. V., 2001, Bayesian Networks and Decision Graphs, Statistics for Engineering and Information Science, Springer.
[25] Microsoft Research Center, online MSBNx Editor Manual and software download, `http://research.microsoft.com/adapt/MSBNx/`.
[26] M. R. Rosekind, E. L. Co, K. B. Gregory, and D. L. Miller, 2000, Crew Factors in Flight Operations XIII: A Survey of Fatigue Factors in Corporate/Executive Aviation Operations, National Aeronautics and Space Administration, Ames Research Center Moffett Field, California 94035, NASA/TM-2000-209610, September.
[27] E. L. Co, K. B. Gregory, J. M. Johnson, and M. R. Rosekind, 1999, Crew Factors in Flight Operations XI: A Survey of Fatigue Factors in Regional Airline Operations, National Aeronautics and Space Administration, Ames Research Center Moffett Field, California 94035, NASA/TM-1999-208799, October.
[28] P. Sherry, 2000, Fatigue Countermeasures in the Railroad Industry-Past and Current

Developments, Counseling Psychology Program, Inter-modal Transportation Institute, University of Denver, June 2000..

[29] M R. Rosekind, K. B. Gregory, E. L. Co, D. L. Miller, and D. F. Dinges, 2000, Crew Factors in Flight Operations XII: A Survey of Sleep Quantity and Quality in On-Board Crew Rest Facilities, National Aeronautics and Space Administration, Ames Research Center Moffett Field, California 94035, NASA/TM-2000-209611, September.

[30] Q. Ji, 2002, 3D Face pose estimation and tracking from a monocular camera. *Image and Vision Computing*, Vol. 20, No. 7, Pp. 499-511.

[31] Q. Ji and Z. Zhu, in press, Non-intrusive monitoring and prediction of driver fatigue. *IEEE Transactions on Vehicular Technology*.

[32] Q. Ji and Z. Zhu, 2002, Eye and gaze tracking for interactive graphic display. *Smart Graphics*. Hawthorne, NY.

[33] Z. Zhu, K. Fujimura, and Q. Ji., 2002, Real-time eye detection and tracking under various light conditions and face orientations. *2002 ACM SIGCHI Symposium on Eye Tracking Research & Applications*. New Orleans, LA.

[34] H. Gu, Q. Ji, and Z. Zhu, 2002, Active facial tracking for fatigue detection. *IEEE Workshop on Applications of Computer Vision*. Orlando, FL.

[35] Q. Ji and X. Yang, 2002, Real time 3D face pose discrimination based on active IR illumination. *International Conference on Pattern Recognition* .

[36] Z. Zhu, Q. Ji, and K. Fujimura, 2002, Combining Kalman filtering and mean shift for real time eye tracking under active IR illumination. *International Conference on Pattern Recognition*

[37] Zhiwei Zhu and Qiang Ji, Real time 3D face pose tracking from an uncalibrated camera, submitted to IEEE transactions on Pattern Analysis and Machine Intelligence.

[38] Yongmian Zhang and Q. Ji, 2003, Expression Understanding in Image Sequences with Dynamic and Active Visual Information Fusion, International Conference on Computer Vision, Nice, France.

[39] Xiangyang Li and Qiang Ji, 2003, Active Affective State Detection and Assistance with Dynamic Bayesian Networks, 3rd Workshop on Affective and Attitude User Modeling Assessing and Adapting to User Attitudes and Affect: Why, When and How? in conjunction with User Modeling 2003 Pittsburgh, PA, USA June 22..

[40] Peilin Lan and Q. Ji, 2003, Non-invasive real time human fatigue modeling and monitoring. The 47th Human Factors and Ergonomics Society Annual Meeting. Denver, Colorado, October 13-17.

[41] P. Lan, Q. Ji, and C. Looney, 2002, Information fusion of visual cues using dynamic Bayesian belief networks for inferring human cognitive states. The Invited Session on Information Fusion using Bayesian Networks (Fusion 2002). Annapolis, MD.

[42] Reiner Onken,1994, DAISY, an adaptive knowledge-based driver monitoring and warning system, Proceedings of 1994 Vehicle navigation and information systems conference, Yokohama, Japan, pages 3-10.

[43] Richard Grace, 1999, A drowsy driver detection system for heavy vehicles", Conference on

ocular measures of driver alertness, April.
[44] Dixon Cleveland, 1999, Unobtrusive eyelid closure and visual of regard measurement system, Conference on ocular measures of driver alertness, April.
[45] J. Fukuda, K. Adachi, M. Nishida and E Akutsu, 1995, Development of driver's drowsiness detection technology, Toyota technical review, Vol. 45, pages 34-40.
[46] J. H. Richardson, 1995, the development of a driver alertness monitoring system, Fatigue and Driving: Driver impairment, driver fatigue and driver simulation, Taylor and Francis; London, L. Hartley (Ed).
[47] P. Ekman and W. V. Friesen, 1978, Facial Action Coding System (FACS): Manual, Consulting Psychologists Press, Palo Alto, CA.